The State of Data in Development Economics: Big, Open and Meta

Tariq Khokhar: Data Scientist, World Bank
@tkb & @worldbankdata
$ + Database = THE WORLD BANK
What do development economists do?
Since the beginning of 2014, the rate of inflation has remained steady at around six percent, largely due to the stability of food and energy prices on international markets. Interest rates also appear to have remained stable (see Figure 10). The value of the local currency relative to the US dollar depreciated marginally in the period from January to November 2014, with the rate of depreciation accelerating towards the end of that period due to a shift in the composition of monetary growth (see Figure 11).

Figure 10 and 11: Stable inflation, exchange and interest rates with depreciation of the local currency value from January to November 2014

Source: World Bank

The drivers of growth are the construction, communication, health, electricity, manufacturing, and other service sectors.
Inflation as measured by the consumer price index reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly. The Laspeyres formula is generally used.

International Monetary Fund, International Financial Statistics and data files.

<table>
<thead>
<tr>
<th>Country name</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>-8.3</td>
<td>0.9</td>
<td>10.2</td>
<td>7.2</td>
</tr>
<tr>
<td>Albania</td>
<td>2.3</td>
<td>3.6</td>
<td>3.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Algeria</td>
<td>5.7</td>
<td>3.9</td>
<td>4.5</td>
<td>8.9</td>
</tr>
<tr>
<td>American Samoa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andorra</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IMF eLibrary Data

Get predefined data reports by country, topic, or data source

Or build your own queries across or within datasets

SIGN IN OR REGISTER ▸
START YOUR 30-DAY FREE TRIAL ▸

Most Visited Data and Statistics

World Economic Outlook April 2014 new
Global Financial Stability Report April 2014 new
Fiscal Monitor April 2014 new

IMF ELIBRARY DATA:
IFS, BOP, DOTS, GFS
IMF Data Mapper®
IMF Finances new
SDRs per Currency Unit

Not sure where to find what you're looking for?
A Guide to IMF Data
Selected IMF Online Data Sources

IMF Data Mapper® 3.0
New Data and Features: African REO now available!

Data and Statistics
Last Updated: April 28, 2014

The IMF publishes a range of time series data on IMF lending, exchange rates and other economic and financial indicators. Manuals, guides, and other material on statistical practices at the IMF, in member countries, and of the statistical community as a whole are also available.
<table>
<thead>
<tr>
<th>IMF Code</th>
<th>ISO Code</th>
<th>Country</th>
<th>Historical Data Source</th>
<th>Latest Actual Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>144</td>
<td>SWE</td>
<td>Sweden</td>
<td>NSO</td>
<td>2013</td>
</tr>
<tr>
<td>146</td>
<td>CHE</td>
<td>Switzerland</td>
<td>NSO</td>
<td>2013</td>
</tr>
<tr>
<td>463</td>
<td>SYR</td>
<td>Syria</td>
<td>NSO</td>
<td>2011</td>
</tr>
<tr>
<td>528</td>
<td>TWN</td>
<td>Taiwan Province of China</td>
<td>NSO</td>
<td>2013</td>
</tr>
<tr>
<td>923</td>
<td>TJK</td>
<td>Tajikistan</td>
<td>NSO</td>
<td>2012</td>
</tr>
<tr>
<td>738</td>
<td>TZA</td>
<td>Tanzania</td>
<td>NSO</td>
<td>2013</td>
</tr>
<tr>
<td>578</td>
<td>THA</td>
<td>Thailand</td>
<td>NSO</td>
<td>2013</td>
</tr>
<tr>
<td>537</td>
<td>TLS</td>
<td>Timor-Leste</td>
<td>NSO</td>
<td>2012</td>
</tr>
<tr>
<td>742</td>
<td>TGO</td>
<td>Togo</td>
<td>NSO</td>
<td>2013</td>
</tr>
<tr>
<td>866</td>
<td>TON</td>
<td>Tonga</td>
<td>CB</td>
<td>2012</td>
</tr>
<tr>
<td>369</td>
<td>TTO</td>
<td>Trinidad and Tobago</td>
<td>NSO</td>
<td>2013</td>
</tr>
<tr>
<td>744</td>
<td>TUN</td>
<td>Tunisia</td>
<td>NSO</td>
<td>2012</td>
</tr>
<tr>
<td>186</td>
<td>TUR</td>
<td>Turkey</td>
<td>NSO</td>
<td>2013</td>
</tr>
<tr>
<td>925</td>
<td>TKM</td>
<td>Turkmenistan</td>
<td>NSO</td>
<td>2012</td>
</tr>
<tr>
<td>869</td>
<td>TUV</td>
<td>Tuvalu</td>
<td>NSO</td>
<td>2012</td>
</tr>
<tr>
<td>746</td>
<td>UGA</td>
<td>Uganda</td>
<td>CB</td>
<td>2013/14</td>
</tr>
</tbody>
</table>
The National Bureau of Statistics (NBS) is the National Statistical Office of Tanzania. NBS, as a Government agency, was established from the former Central Bureau of Statistics to provide efficient services, giving increased value for money for the benefit of the Government and the public in general. NBS carries out its activities in a businesslike manner, using commercial financial management and business-planning techniques and is customer focused. It conducts Censuses and Surveys which yield a wide range of economic, social, and demographic statistics.

Click here for The Statistics Act, 2002
PRESS RELEASE

Date: 08th April, 2014

NATIONAL CONSUMER PRICE INDEX (NCPI) FOR MARCH, 2014

Background notes

Definition

The NCPI measures the change over time in the cost of a fixed basket of goods and services that are purchased by a representative sample of households in Tanzania. A basket includes a list of 224 goods and services that are being priced on monthly basis.

Weights and a reference period
ANNUAL HEADLINE INFLATION RATE SLIGHTLY INCREASES TO 6.1%

Table 1: Changes in the NCPI for March, 2014 (September 2010 = 100)

<table>
<thead>
<tr>
<th>S/N</th>
<th>Main Groups</th>
<th>Weight</th>
<th>March-13</th>
<th>Feb-14</th>
<th>March-14</th>
<th>1 Month Percent Change</th>
<th>12 Month Percent Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Food and Non Alcoholic Beverages</td>
<td>47.8</td>
<td>154.95</td>
<td>164.83</td>
<td>166.14</td>
<td>0.8</td>
<td>7.2</td>
</tr>
<tr>
<td>2</td>
<td>Alcoholic and Tobacco</td>
<td>3.3</td>
<td>132.09</td>
<td>141.43</td>
<td>141.47</td>
<td>0.0</td>
<td>7.1</td>
</tr>
<tr>
<td>3</td>
<td>Clothing and Footwear</td>
<td>6.7</td>
<td>127.79</td>
<td>132.00</td>
<td>131.99</td>
<td>0.0</td>
<td>3.3</td>
</tr>
<tr>
<td>4</td>
<td>Housing, Water, Electricity, Gas and Other Fuel</td>
<td>9.2</td>
<td>152.13</td>
<td>164.49</td>
<td>165.85</td>
<td>0.8</td>
<td>9.0</td>
</tr>
<tr>
<td>5</td>
<td>Furnishing, Housing Equipment and Routine Maintenance of the House</td>
<td>6.7</td>
<td>126.84</td>
<td>128.82</td>
<td>129.12</td>
<td>0.2</td>
<td>1.8</td>
</tr>
<tr>
<td>6</td>
<td>Health</td>
<td>0.9</td>
<td>108.08</td>
<td>109.11</td>
<td>110.33</td>
<td>1.1</td>
<td>2.1</td>
</tr>
<tr>
<td>7</td>
<td>Transport</td>
<td>9.5</td>
<td>115.78</td>
<td>125.86</td>
<td>125.84</td>
<td>0.0</td>
<td>8.7</td>
</tr>
<tr>
<td>8</td>
<td>Communication</td>
<td>2.1</td>
<td>96.24</td>
<td>96.67</td>
<td>96.71</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td>9</td>
<td>Recreation and Culture</td>
<td>1.3</td>
<td>111.87</td>
<td>113.15</td>
<td>113.31</td>
<td>0.1</td>
<td>1.3</td>
</tr>
<tr>
<td>10</td>
<td>Education</td>
<td>1.7</td>
<td>113.16</td>
<td>119.16</td>
<td>119.17</td>
<td>0.0</td>
<td>5.3</td>
</tr>
<tr>
<td>11</td>
<td>Restaurants and hotels</td>
<td>6.4</td>
<td>133.52</td>
<td>134.94</td>
<td>135.78</td>
<td>0.6</td>
<td>1.7</td>
</tr>
<tr>
<td>12</td>
<td>Miscellaneous goods and services</td>
<td>4.5</td>
<td>118.77</td>
<td>124.28</td>
<td>125.00</td>
<td>0.6</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>TOTAL – ALL ITEMS INDEX</td>
<td>100</td>
<td>140.93</td>
<td>148.62</td>
<td>149.49</td>
<td>0.6</td>
<td>6.1</td>
</tr>
<tr>
<td>S/N</td>
<td>MAJOR GROUPS</td>
<td>Weights</td>
<td>Jan-14</td>
<td>Feb-14</td>
<td>Mar-14</td>
<td>Apr-14</td>
<td>May-14</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------------------------</td>
<td>---------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>1</td>
<td>Food and Non Alcoholic Beverages</td>
<td>47.8</td>
<td>161.02</td>
<td>164.83</td>
<td>166.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Alcoholic, Tobacco and Narcotics</td>
<td>3.3</td>
<td>141.36</td>
<td>141.43</td>
<td>141.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Clothing and Footwear</td>
<td>6.7</td>
<td>131.76</td>
<td>132.00</td>
<td>131.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Housing, Water, Electricity, Gas and Other Fuel</td>
<td>9.2</td>
<td>163.49</td>
<td>164.49</td>
<td>165.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Furnishing, Housing Equipment and Routine Maintenance of the House</td>
<td>6.7</td>
<td>128.60</td>
<td>128.82</td>
<td>129.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Health</td>
<td>0.9</td>
<td>108.75</td>
<td>109.11</td>
<td>110.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Transport</td>
<td>9.5</td>
<td>125.60</td>
<td>125.86</td>
<td>125.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Communication</td>
<td>2.1</td>
<td>96.73</td>
<td>96.67</td>
<td>96.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Recreation and Culture</td>
<td>1.3</td>
<td>113.01</td>
<td>113.15</td>
<td>113.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Education</td>
<td>1.7</td>
<td>119.14</td>
<td>119.16</td>
<td>119.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Restaurants and hotels</td>
<td>6.4</td>
<td>134.85</td>
<td>134.94</td>
<td>135.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Miscellaneous goods and services</td>
<td>4.5</td>
<td>123.54</td>
<td>124.28</td>
<td>125.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td><strong>TOTAL – ALL ITEMS INDEX</strong></td>
<td><strong>100.0</strong></td>
<td><strong>146.60</strong></td>
<td><strong>148.62</strong></td>
<td><strong>149.49</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**INFLATION RATES**

<table>
<thead>
<tr>
<th>Other Selected Groups</th>
<th>Weight</th>
<th>Jan-14</th>
<th>Feb-14</th>
<th>Mar-14</th>
<th>Apr-14</th>
<th>May-14</th>
<th>Jun-14</th>
<th>Jul-14</th>
<th>Aug-14</th>
<th>Sep-14</th>
<th>Oct-14</th>
<th>Nov-14</th>
<th>Dec-14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Food - combining food consumed at home and food consumed in restaurants</td>
<td>51.0</td>
<td>159.69</td>
<td>163.29</td>
<td>164.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Energy Fuels - combining electricity and other fuels for use at home with petrol and diesel</td>
<td>5.7</td>
<td>184.25</td>
<td>185.08</td>
<td>187.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 All Items Less Food</td>
<td>49.0</td>
<td>132.96</td>
<td>133.32</td>
<td>133.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 All Items Less Food and Energy</td>
<td>43.3</td>
<td>125.88</td>
<td>126.17</td>
<td>126.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Food Inflation Rate                                      | 6.6    | 6.9    | 7.9    |        |        |        |        |        |        |        |        |        |        |
| Energy Inflation Rate                                    | 18.7   | 16.1   | 10.3   |        |        |        |        |        |        |        |        |        |        |
| Non Food Inflation Rate                                  | 6.7    | 6.3    | 5.2    |        |        |        |        |        |        |        |        |        |        |
| Non Food and non Energy Inflation Rate                   | 4.9    | 4.8    | 4.7    |        |        |        |        |        |        |        |        |        |        |
Tanzania
Inflation, consumer prices (annual %)

Inflation as measured by the consumer price index reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly. The Laspeyres formula is generally used.

International Monetary Fund, International Financial Statistics and data files.

Catalog Sources: World Development Indicators
Open Data
From this

World Development Indicators 2009

by: World Bank

Price: $75.00  *Geographic discounts available!

Available

World Development Indicators
English; Paperback; 456 pages; 8.5x11
Published April 24, 2009 by World Bank
ISBN: 978-0-8213-7829-8; SKU: 17829
From this
World Bank Open Data: free and open access to data about development in countries around the globe.

ICP 2011: 7 million prices, 199 economies, 8 regions, and 15 partners
Haishan Fu | Posted on 30 Apr 2014

Updates to WDI 2014: Purchasing Power Parities and Maternal Mortality Ratios
Posted on 2 May 2014

WITS trade data site: Five new features
Siddheesh Kaushik | Posted on 22 Apr 2014

Find an indicator
GNI per capita, Atlas method (current US$) Go

BROWSE DATA

FEATURED
World Development Indicators
Open Finances
Projects & Operations
Open Government Data Toolkit

RECENTLY UPDATED
Gender Statistics 16 Apr, 2014
World Development Indicators 09 Apr, 2014

The World at a Glance
Economy & Growth
GDP (current US$)
$72.49 trillion 2012

Health
Life expectancy at birth, total (years)
71 2012

For developers Terms of use Email updates
India

South Asia

GDP (current US$)
$1.859 trillion 2012

Income level
Lower middle income
Population (Total)
1.237 billion 2012

World Development Indicators

Graph, map and compare more than 1,000 time series indicators from the World Development Indicators and more than 5,000 indicators from other collections such as Gender Statistics, African Development Indicators, and Education Statistics.

School enrollment, primary (% gross)

113% 2011

CO2 emissions (metric tons per capita)

1.7 2010
Global Economic Prospects - Forecasts

Global Economic Prospects examines trends for the world economy and how they affect developing countries. The report includes country-specific three-year forecasts for major macroeconomic indicators, including commodity and financial markets.

Projects & Operations

Search, browse and map more than 10,000 projects from 1947 to the present.
India

South Asia

GDP (current US$) $1.859 trillion 2012

Income level

Lower middle income 1.237 billion 2012

World Development Indicators

Graph, map and compare more than 1,000 time series indicators from the World Development Indicators and more than 5,000 indicators from other collections such as Gender Statistics, African Development Indicators, and Education Statistics.

DATABANK

DOWNLOAD DATA

School enrollment, primary (% gross)

113% 2011

CO2 emissions (metric tons per capita)

1.7 2010
भारत

दक्षिण एशिया

$1.859 10 खर्च 2012

आय स्रोत

निम्न-मध्यम आय

1.237 1 अरब 2012

विश्व विकास सूचक

विश्व विकास के सूचकों के 1,000 से अधिक समय बंदरों के सूचकों और 5,000 से अधिक अतिरिक्त सूचकों के ग्राफ बनाएं, खाना उत्पादन करें (मैदा) और तुलना करें जो अन्य उद्योग के जैसे वैश्विक संख्याकीय, आँकड़ों के विकास सूचकों या विकास सांख्यिकी में निहित है।
|--------------|--------------|----------------|----------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
Data Catalog

The World Bank’s Open Data initiative is intended to provide all users with access to World Bank data, according to the Open Data Terms of Use. The data catalog is a listing of available World Bank datasets, including databases, pre-formatted tables, reports, and other resources.

163 records found

Sort By: Most Popular | Alphabetical | Last Updated

World Development Indicators (WDI) | New

The primary World Bank collection of development indicators, compiled from officially-recognized international sources. It presents the most current and accurate global development data available, and includes national, regional and global estimates.

Type: Time series | Periodicity: Annual | Last Updated: 09-Apr-2014

GDP ranking (GDP) | New

Gross domestic product (GDP) ranking.

Type: Time series | Periodicity: Annual | Last Updated: 09-Apr-2014
Central Microdata Catalog

Found 1591 studies out of 1591

Sort results by: Country ▲ | Year | Title | Popularity | Collection

Showing 1-15 of 1591 studies

Enterprise Survey 2013
Afghanistan, 2013
By: World Bank
Collection: Enterprise Surveys

World Bank Country Survey 2012
Afghanistan, 2012
By: Public Opinion Organization
Collection: The World Bank Group Country Opinion Survey Program (COS)
What have we learned?
Inflation, consumer prices (annual %) for Tanzania

- Year 2009: 12.1422278742992
- Year 2008: 10.2783937621131
- Year 2007: 7.02551436974841
- Year 2006: 7.250972620825
Google “gdp Tanzania”
Polio in Africa: 1988 vs 2012

via Voice of America with data from the CDC and GPEI

Afghanistan, Pakistan and Nigeria are the only three countries where polio is still endemic.
The distribution of poverty is changing: who's living on $1.25 a day?

2021
Number of people worldwide living in extreme poverty, in millions

<table>
<thead>
<tr>
<th>Worst case</th>
<th>Baseline</th>
<th>Best case</th>
</tr>
</thead>
<tbody>
<tr>
<td>1084.8</td>
<td>620.3</td>
<td>325.6</td>
</tr>
</tbody>
</table>

Since 1990, more than 700 million people have risen from extreme poverty (living on US$1.25 a day or less), but more than 1 billion people remain extremely poor today. And 1 in 3 of them are children.
(Often) The Present
There are gaps & lags in the data
Big Data
# Inflation, consumer prices (annual %)

Inflation as measured by the consumer price index reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly. The Laspeyres formula is generally used.

International Monetary Fund, International Financial Statistics and data files.

**Catalog Sources** World Development Indicators

<table>
<thead>
<tr>
<th>Country name</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>-8.3</td>
<td>0.9</td>
<td>10.2</td>
<td>7.2</td>
</tr>
<tr>
<td>Albania</td>
<td>2.3</td>
<td>3.6</td>
<td>3.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Algeria</td>
<td>5.7</td>
<td>3.9</td>
<td>4.5</td>
<td>8.9</td>
</tr>
<tr>
<td>American Samoa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andorra</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Tanzania

http://bit.ly/1fJRvEC
PriceStats uses online prices to calculate daily inflation statistics.

1. Scraping
2. Online Retailers
3. Processing
4. Inflation Statistics
US INFLATION SERIES
PriceStats estimates aggregate inflation in the US using online prices. The objective of this series is to anticipate major changes in US inflation trends, but not to forecast monthly CPI announcements. At any point in time, our index can be different from the CPI. Our data anticipates changes in inflation trends not only because we observe prices sooner, but also because online prices tend to react to shocks more quickly.

- Official CPI
- PriceStats Index

Visit State Street Research Portal to download the data.

http://www.pricestats.com/us-series
<table>
<thead>
<tr>
<th>Country</th>
<th>Product</th>
<th>Index</th>
<th>30 Day</th>
<th>7 Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>Staples</td>
<td>101.6</td>
<td>-0.13%</td>
<td>-0.09%</td>
</tr>
<tr>
<td>Argentina</td>
<td>Bread</td>
<td>117.5</td>
<td>-3.32%</td>
<td>-0.8%</td>
</tr>
<tr>
<td>India</td>
<td>Onions</td>
<td>210.9</td>
<td>-8.73%</td>
<td>-6.46%</td>
</tr>
</tbody>
</table>

View a real-time monitor of consumer staples prices in China. Click components to access thousands of granular daily observations. Explore.

View a real-time map of consumer staples, for example bread, in emerging markets. Explore.

Monitor price, quantity and availability of this critical staple ingredient across marketplaces in Mumbai, Chennai, and Kolkata. Explore.
Call Detail records (CDRs) are metadata that capture subscribers’ use of their cell-phones — including an identification code and, at a minimum, the location of the phone tower that routed the call for both caller and receiver — and the time and duration of call. Large operators collect over six billion CDRs per day.

<table>
<thead>
<tr>
<th>CALLER ID</th>
<th>CALLER CELL TOWER LOCATION</th>
<th>RECIPIENT PHONE NUMBER</th>
<th>RECIPIENT CELL TOWER LOCATION</th>
<th>CALL TIME</th>
<th>CALL DURATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>X76VG588RLPQ</td>
<td>2°24' 22.14&quot;, 35°49' 56.54&quot;</td>
<td>A81UTC93KK52</td>
<td>3°26' 30.47&quot;, 31°12' 18.01&quot;</td>
<td>2013-11-07T15:15:00</td>
<td>01:12:02</td>
</tr>
</tbody>
</table>

http://www.unalobalpulse.org/Mobile Phone Network Data-for-Dev
Figure 1: Côte d’Ivoire - Location of Voronoi cells, cell phone antennas (light circles) and DHS clusters (dark circles)
Figure 4: Relationship between the logarithm of population size and the logarithm of Orange mobile phone users at the “sous-prefecture” level in the Côte d’Ivoire.
Figure 5: Observed population density (left) vs predicted population density based on subscriber data (right)
Meta Data
US stocks end lower on GDP revisions

(Close): US markets closed lower on Friday after GDP was revised down to an annualised 2.2% for the last three months of 2014.

However, all three indexes were up overall for February, with the S&P 500 and Dow Jones hitting fresh highs.

The Dow Jones shed 81.72 points to 18,132.70.
Another great quote from #BTN2014: "Data citation. It would be awesome if there were some."
WHAT CAN DATACITE DO FOR YOU?

CITE YOUR DATA
Data citation is fundamental as it enables easy reuse and verification of data, making it possible to track and quantify the impact of data. Citation creates a...

FORMAT YOUR CITATION
Use the DOI Citation Formatter, a service created in collaboration with CrossRef, to format your citation, ensuring you adopt the correct format for your needs...

FIND A REPOSITORY
DataCite supports all researchers looking to deposit and/or find data, in collaboration with the Databib initiative and re3data.org.

FIND A DATASET
DataCite Metadata Search is a service that allows people to search for datasets registered with DataCite, via the metadata associated with the datasets.

GET YOUR DOI STATISTICS
DataCite provides statistics for members on DOI registrations and DOI resolutions, filtered by Allocator, Data Center or Prefix.
Irino, T; Tada, R (2009): Chemical and mineral compositions of sediments from ODP Site 127-797. Geological Institute, University of Tokyo: http://dx.doi.org/10.1594/PANGAEA.726855
The Bechdel Test Over Time
How women are represented in movies

- Fewer than two women
- Women don’t talk to each other
- Women only talk about men
- Dubious
- Passes Bechdel Test

1970-’74
1980-’84
1990-’94
2000-’04
2010-’13

0%
25%
50%
75%
100%
The Need for Openness in Data Journalism

Brian Keegan, Ph.D. (@bkeegan) College of Humanities and Social Sciences, Northeastern University

Do films that pass the Bechdel Test make more money for their producers? I've replicated Walt Hickey's recent article in FiveThirtyEight to find out. My results confirm his own in part, but also find notable differences that point the need for clarification at a minimum. While I am far from the first to make this argument, this case is illustrative of a larger need for journalism and other data-driven enterprises to borrow from hard-won scientific practices of sharing data and code as well as supporting the review and revision of findings. This admittedly lengthy post is a critique of not only this particular case but also an attempt to work through what open data journalism could look like.

The Angle: Data Journalism should emulate the openness of science

New data-driven journalists such as FiveThirtyEight have faced criticism from many quarters and the critiques, particularly around the naïveté of assuming credentialed experts can be bowled over by quantitative analysis so easily as the terrifyingly innumerate pundits who infest our political media. While I find these critiques persuasive, I depart from them here to instead argue that I have found this "new" brand of data journalism disappointing foremost because it wants to perform science without abiding by scientific norms.

The questions of demarcating what is or is not science are fraught, so let's instead label my gripe a "failure to be open." By openness, I don't mean users commenting on articles or publishing whistleblowers' documents. I mean "openness" more in the sense of "open source software" where the code is made freely available to everyone to inspect, copy, modify, and redistribute. But the principles of open-source software trace their roots more directly back norms in the scientific community that Robert Merton identified and came to known as "CUDOS" norms. It's worth reviewing two of these norms because Punk-ass Data Journalism is very much on the lawn of Old Man Science and therein lie exciting possibilities for exciting adventures.
Bechdel Test data from BechdelTest.com's API

Documentation here: http://bechdeltest.com/api/v1/doc#getMovieByImdbId

First, download a list of all movie IDs using the `getAllMovieIds` method. Inspect the first 5. Create a list of IMDB ids out of this list of dictionaries to pass to the API.

```
In [5]:
movie_ids = json.loads(urllib2.urlopen('http://bechdeltest.com/api/v1/getAllMovieIds').read())
imdb_ids = [movie['imdbid'] for movie in movie_ids]
print u"There are {0} movies in the Bechdel test corpus".format(len(imdb_ids))

There are 5062 movies in the Bechdel test corpus
```

Pull an example `getMovieByImdbId` from the API.

```
In [5]:
json.loads(urllib2.urlopen('http://bechdeltest.com/api/v1/getMovieByImdbId?imdbid=000091').read())
```

```
Out[5]:
{u'date': u'2013-12-25 14:31:21',
 u'dubious': u'0',
 u'id': u'4982',
 u'imdbid': u'0000091',
 u'rating': u'0',
 u'submitterid': u'9025',
 u'title': u'House of the Devil, The',
 u'visible': u'1',
 u'year': u'1896'}
```

Pull all of the data we can down. This will take a very long time. We don't want to have to repeat that crawl again, so let's write the data to disk as `bechdel.json`.

```
In []:
ratings = dict()
exceptions = list()
for num,imdb_id in enumerate(imdb_ids):
    try:
        if num in range(0,len(imdb_ids),int(round(len(imdb_ids)/10,0))):
            print imdb_id
        ratings[imdb_id] = json.loads(urllib2.urlopen('http://bechdeltest.com/api/v1/getMovieByImdbId?imdbid={0}'.format(imdb_id)).read())
    except:
        exceptions.append(imdb_id)
pass
```
FiveThirtyEight

Data

Data and code behind the stories and interactives at FiveThirtyEight

Updated an hour ago

app-template

Forked from nprapps/app-template

The NPR visuals team's opinionated project template for client-side apps.

Updated on Feb 5
Old Approach

New Approach

https://karthik.github.io/testdat-talk/
Thanks!

Tariq Khokhar: Data Scientist, World Bank
@tkb & @worldbankdata